DOCUMENT RESUME

ED 383 766 TM 023 312

AUTHOR Weigle, David C.; Snow, Alicia

TITLE Interpreting Where Detected Effects Originate: Structure Coefficients Versus Pattern and Other Coefficients.

PUB DATE 21 Apr 95

NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).

PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Comparative Analysis; *Factor Analysis; Factor Structure; Literature Reviews; *Matrices; *Scores

IDENTIFIERS *Pattern Analysis; *Structure Coefficients

ABSTRACT

Various analytic choices in principal components and common factor analysis are discussed. Differences and similarities among these extraction methods are explained, and aids in interpreting the origin of detected effects are explored. Specifically, the nature and use of structure and pattern coefficients are examined. Communalities and methods for obtaining factor scores are presented, and a selected review of studies pertaining to these topics is also included. (Contains 13 references and 4 tables.) (Author)

Interpreting Where Detected Effects Originate:

Structure Coefficients Versus Pattern and Other Coefficients

David C. Weigle Alicia Snow

Texas A&M University 77843-4225

Paper presented at the annual meeting of the American Educational Research Association (Session #54.21), San Francisco, California, April 21, 1995.

Abstract

Various analytic choices in principal components and common factor analysis are discussed. Differences and similarities among these extraction methods are explained, and aids in interpreting the origin of detected effects are explored. Specifically, the nature and use of structure and pattern coefficients are examined. Communalities and methods for obtaining factor scores are presented, and a selected review of published factor analytic studies pertaining to these topics is also included.

Factor analysis is an analytic technique that permits the reduction of a large number of interrelated variables to a smaller number of latent dimensions (Tinsley & Tinsley, 1987). The goal of factor analysis is to explain the maximum amount of variance by using the smallest number of concepts and thereby provide a meaningful organizational scheme that can be used to interpret the data being analyzed with the greatest parsimony.

Although factor analysis was conceptualized in the early years of this century, it has only come into widespread use with the advent of modern computer technologies. Computerized statistical packages such as SPSS and SAS have made factor analytic techniques readily available to social scientists and educators, but these computerized resources do not relieve researchers from the obligation to make informed analytic decisions. Rather, easy accessibility to rich and complex analyses *demands* that investigators acquire a thorough conceptual understanding of statistical methodologies so that the data under consideration will be interpreted insightfully and meaningfully.

This paper will explore the conceptual bases of certain elements of factor analytic methods relating to specific analytic decisions and the origins of detected effects. Differences and similarities among extraction methods will be explained, and the nature and use of structure and pattern coefficients will be examined. Communalities and methods for obtaining factor scores will be presented, and a selected review of published factor analytic studies pertaining to these topics will be included. It is expected that when armed with conceptual understanding, even the novice researcher will become a more informed and effective consumer of the professional literature.

Principal Components vs. Common Factor Analyses

Principal components and common or principal factor analysis are the primary methods of exploratory factor analysis. A matrix of associations--

typically a correlation matrix--forms the basis of the analysis for either of these extraction methods. The difference between the two approaches involves the entries used on the diagonals of the matrix being analyzed.

Principal components uses ones (1.000s) on the diagonals as reliability estimates. Within this method, therefore, perfect reliability is assumed and all variance can be accounted for. Variance can be common to the factor or specific to the variable. The resulting reproduced correlation matrix is the best least-squares estimate of the entire correlation matrix, including the diagonal elements of the correlation matrix. In other words, principal components results in the sum-of-the-squared differences between the original and reproduced correlation matrices being minimized.

The same analytic process is used for common factor analysis, but the initial matrix is altered so that estimates of reliabilities, rather than ones, are used on the diagonals. Then iteration is employed to further refine these initial estimates. The assumption of less than perfect reliability allows for error and the possibility that the variables will not be perfectly reproduced by the common factors alone. Common factor methods, therefore, look at common variance rather than total variance.

A variety of common factor methods exist (*e.g.*, image analysis, alpha factor analysis and principal axis analysis). The differences between the common factor methods lie primarily in the alterations to the correlation matrix before the final factors are extracted. The most widely used variant of common factor analysis is principal axis analysis in which communalities (measures of variance accounted for by the common factor) are placed in the diagonals.

Principal components and common factor analysis yield increasingly more equivalent results as either (a) the factored variables are more reliable or (b) the number of variables being factored is increased. Snook and Gorsuch (1989, p. 149) explain this second point noting that, "As the number of variables decreases, the ratio of diagonal to off-diagonal elements also decreases, and therefore, the value of the communality has an increasing effect on the analysis." For example, with 10 variables, the 10 diagonal entries represent 10% (10/100) of the one hundred entries in the matrix. With 100 entries, however, the diagonal entries represent only 1% (100/10,000) of the 10,000 matrix entries.

Analysts differ quite heatedly over the utility of principal components as compared to common or principal factor analysis. An entire special issue of Multivariate Behavioral Research was devoted to this controversy. Gorsuch (1983) prefers common factor to principal components because (a) few variables are thought to be error free, (b) common factor produces more conservative loadings (lower factor structure coefficients), and (c) common factor produces a principal components analysis if it is truly appropriate. Velicer and Jackson (1990) hold that the two methods are similar and only yield discrepant results when too many factors are extracted. They conclude that both methods are equally generalizable while Widaman (1993) concluded "that the parameters defined by CFA are more generalizable than those defined by component analysis."

Structure and Pattern Coefficients

Two matrices are derived to explain the relationship between the extracted factors and the original variables. Elements of the factor pattern matrix represent standardized linear weights analogous to beta weights in regression (Thompson & Borrello, 1985) and are known as pattern coefficients. Pattern coefficients indicate the importance of a given variable to the factor with the influence of the other variables partialled out.

Elements of the factor structure matrix are simple correlations between variables and the factor scores or latent variable composites that are derived through a process of weighting and aggregating using the pattern coefficients. These structure coefficients are sometimes called factor loadings. When the factors are perfectly uncorrelated the pattern and structure matrices are the same. It follows that these matrices are identical when orthogonal rotation procedures are used, but they are different when oblique rotation procedures are used since oblique rotation results in correlated factors.

Table 1 presents the items that comprise a newly developed scale of aggression that has been created for inclusion in the Marital Satisfaction Inventory (Snyder & Snow, 1994). The pattern and structure coefficients

---

Insert Table 1 about here

---

obtained from a principal components analysis with a varimax (orthogonal) rotation of data derived from this scale are presented in Table 2. Note that the pattern and structure coefficients are the same in this case.

Insert Table 2 about here

In Table 3, however, the coefficients are no longer equal. The coefficients in Table 3 represent the results of a component analysis with an oblimin (oblique) rotation that yields correlated factors. In this instance a .51 correlation exists between Factor I and Factor II.

Insert Table 3 about here

Because the pattern and structure matrices are identical when the factors are uncorrelated, the question of matrix interpretation becomes more complex only when the factors are correlated. Conclusions can be drawn about the nature of a factor by examining the magnitude of the structure coefficients. Examining structure coefficients (or factor loadings) can aid in naming the emerging factors.

Interpreting the structure matrix as opposed to other matrices--including the factor pattern matrix--has some advantages (Gorsuch, 1974). Investigators are practiced at interpreting correlation coefficients, and the structure matrix presents the simple (zero order) relationship between the variable and the factor. The pattern matrix and reference vector correlations exclude overlap among the factors and represent only unique contributions, even when the overlap is important (Gorsuch, 1983). Structure coefficients also represent a stable relationship between a variable and a factor across studies. Pattern coefficients and reference vectors can only be interpreted within the context of a given study, because the pattern structure shifts if either the factors or the variables in the study are changed.

Pattern coefficients do have some uses. For example, the factor pattern matrix can also be used to reproduce the correlations between variables in order to evaluate the adequacy of a solution (Harman, 1967). The reproduced portion of the correlation matrix equals $P_{VxF} P'_{FxV}$.

Both structure and factor pattern coefficients must be used to gain a clear picture of the interrelationships between the variables (Comrey, 1973) when an oblique rotation is performed. Although structure coefficients are commonly examined, pattern coefficients and reference vector correlations (part correlations between a variable and a factor when the variance attributable to all other factors has been partialled out) should not be dismissed because they contribute valuable information about the unique contribution of each factor to each variable. Ideally, each of these matrices should be included in research reports to provide readers with more information upon which to base conclusions (Gorsuch, 1983).

Communality Coefficients

The proportion of the variance of each variable that is reproduced by the extracted factors is called" communality" and can be considered a lower bound estimate of the reliability of variables. Indeed, in principal axis analysis, communalities are used on the diagonals as reliability estimates. In the case of uncorrelated principal components the communality is derived by summing together all the squared structure coefficients for a variable after the components are extracted.

The proportion of variance accounted for by the factor structure can be determined by dividing the sum of the communalities by the number of factors. Communality coefficients in factor analysis can play an important role in interpretation. For example, a small (*e.g.*, $< |.30|$) final communality for a variable would indicate that little of the variable's variance is explained by the factor structure. In such a case the researcher might consider dropping the variable from the analysis or, alternatively, extracting more factors.

Factor Scores

A factor score is a latent variable consisting of a weighted combination of the scores on each of the variables (Kachigan, 1982, p. 244). Factor scores are hypothetical constructs that represent individuals' scores on factors as opposed to their scores on variables. Computation of factor scores limit the need to conduct a formal factor analysis in future research and provide for the continuing investigation of constructs. They also make it possible to relate the factors to variables (such as nominal variables) that cannot be related by any other analysis. In the case of standardized, noncentered factor scores (Thompson, 1983, 1993), they provide for meaningful comparisons of mean differences across factor scores.

A variety of methods are commonly used to obtain factor scores, but the best procedures share the following characteristics: (a) scores should have high correlations with the factors they are measuring, (b) scores should be unbiased estimates of true factor scores, (c) scores should be univocal (have zero correlation with other factors), (d) orthogonal factor scores should not correlate with each other, and (e) the correlations among correlated factor scores should equal the correlations among the factors. For example, if Factor A correlates .4 with B, then the factor scores on A should correlate .4 with the factor scores on B.

The regression method for obtaining factor scores is popular because of widespread familiarity with regression analyses and because of the fact that this method is the default method in major statistical packages such as SPSS and SAS. This method uses least-squares regression logic to calculate factor scores and arrives at a least squares solution such that the correlation between the underlying factor and the factor scores is maximized. Unfortunately, regression estimates of common factor scores are neither univocal nor unbiased.

Estimates of factor scores obtained through the idealized variables method are based on ideal variables and assume that the variables are perfectly measured. This is a least squares method that differs from the regression method in that it uses the reproduced correlation matrix rather than the original correlation matrix to compute the factor scores. Because the idealized variables matrix and the observed variables matrix are identical in principal components, the two methods will yield identical results. However, the common factor estimates will yield increasingly dissimilar results as the sample correlations diverge from the population correlations. Factor scores obtained by this method are considered unbiased and univocal, but they may correlate somewhat even when factors are orthogonal (Gorsuch, 1983).

Bartlett's method is also commonly used to obtain factor scores. This method provides a least squares procedure designed to minimize the sum of squares of the unique factors over the range of the variables. It gives less weight to those variables with high uniqueness (low communality coefficients) and more weight to variables with low uniqueness. Because error is not considered in principal components analysis, this method results in the same factor scores as the previous methods. Bartlett's method is considered the best procedure for producing univocal factor scores (Kim & Meuller, 1987).

Anderson and Rubin's method for computing factor scores is nearly identical to Bartlett's method, but with the additional requirements that the factor scores be orthogonal. This method produces factor score estimates that are neither univocal nor unbiased.

Each of these methods for computing factor scores produces scores with a mean of zero and a standard deviation of one on each factor. It is impossible, therefore, to make meaningful comparisons of mean differences across factors. Thompson (1983, 1993) describes an alternative method for obtaining factor scores whereby the raw scores are divided by the standard deviation, thus producing standardized, non-centered factor scores. These scores retain the standard deviation of 1.0 and the correlation coefficients match those among the principal components, but the means are no longer zero. This method allows for comparison of mean differences in factor scores.

Selected Review of Published Factor Analytic Studies

An often-voiced criticism of published factor-analytic studies is that researchers often fail to include information needed for an appropriate interpretation of the results. For example, failure to provide information regarding how the number of factors was determined or which method of rotation was used can make it difficult to replicate a study. A small scale review of studies that used factor-analytic methods and that were published in the Journal of Consulting and Clinical Psychology (JCCP) was, therefore, conducted to explore recent practice in this area. Seven articles published in 1980 and 16 articles published from 1990-1994 in JCCP were examined for factor extraction method, rotation, and provision of relevant matrices. The results of the review are found in Table 4.

Insert Table 4 about here

In the articles published since 1990, only 6 of 16 articles indicated how the number of factors was chosen. Five of the seven articles published in 1980 specified that they used the "eigenvalue greater than one" rule. Of the studies published since 1990, 75% (12 studies) used principal component analysis and the remainder did not specify the extraction method used. Sixty-three percent of the studies used varimax rotation, 19% stated that an orthogonal rotation procedure was used, and 12% did not specify a rotation method. All the articles published in 1980 specified the use of varimax rotation, and of these studies, 2 (29%) used common factor analysis, 2 used principal component analysis, and 3 (42%) did not specify.

In summary, it appears that principal components analysis has been favored over common factor analysis by authors who publish factor-analytic studies in JCCP. Orthogonal rotation procedures were more common that oblique rotation procedures, and one might speculate that among the studies that did not specify the type of rotation used, most researchers probably used varimax or another type of orthogonal rotation. These trends might indicate that researchers are relying on default settings of the statistical computer packages rather than making informed analytic decisions based on substantive or theoretical criteria.

It was, unfortunately, not uncommon for authors to omit the pattern and structure matrices from their publications. Some authors reported only the more salient factor loadings (structure coefficients). The declining amount of information provided in more recent studies might be attributed to the fact that factor analysis was not the primary analytic focus in many of the studies. Ninety-four percent of the cited studies published from 1990-1994 included a factor analysis that was either preliminary or secondary in importance to the major statistical analysis in the study. Only 43% of the cited articles from 1980 included factor analyses that were either secondary or preliminary to other statistical procedures.

Perhaps some vital information pertaining to factor analysis is omitted from these articles by space-conscious editors. Comrey (1973) suggests that necessary matrices be made available by request or through auxiliary publication sources. Only three of the seven articles from 1980 and five of the sixteen more recent articles invited interested individuals to request further information.

Conclusion

The present paper has explored the conceptual bases of certain elements of factor analytic methods relating to specific analytic decisions and the origins of detected effects. Differences and similarities among extraction methods were examined, and the nature and use of structure and pattern coefficients was discussed. Communalities and methods for obtaining factor scores were presented, and a selected review of published factor analytic studies pertaining to these topics presented. It is hoped that this review of some of the fundamental concepts of factor analysis will inspire readers to become critical consumers of the professional literature and more informed and creative researchers.

References

Comrey, R. L. (1973). First course in factor analysis. New York: Academic Press.

Gorsuch, R. L. (1974). Factor analysis. Philadelphia: W. B. Saunders Co.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Harman, H. H. (1967). Modern factor analysis. Chicago: University of Chicago Press.

Kachigan, S. K. (1982). Multivariate statistical analysis. New York: Radius Press.

Kim, J., & Meuller, C. W. (1987). Factor analysis: Statistical methods and practical issues. Newbury Park, CA: Sage Publications.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. Psychological Bulletin, 106, 148-154.

Snyder, D. K., & Snow, A. C. (1994, November). Measuring spouse aggression. Paper presented at the annual meeting of the Texas Psychological Association, Houston, Texas.

Thompson, B. (1983, January). The calculation of factor scores: An alternative. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Thompson, B. (1993). Calculation of standardized, non-centered factor scores: An alternative to conventional factor scores. Perceptual and Motor Skills, 77, 1128-1130.

Thompson, B. & Borrello, G. M. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. Journal of Counseling Psychology, 34, 414-424.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some further observations. Multivariate Behavioral Research, 25, 97-114.

Table 1

Items on Marital Satisfaction Inventory's Aggression Scale

1. My partner has slammed things around or thrown things in anger.
2. My partner has held me against my will.
3. My partner tries to use his or her anger to control me.
4. My partner has never pushed me or grabbed me in anger.
5. My partner does not insult me or try to humiliate me when he or she is angry.
6. My partner has never thrown things at me in anger.
7. My partner sometimes tries to make me feel inadequate or inferior.
8. My partner has slapped me.
9. My partner does not become verbally abusive when he or she is angry.
10. My partner has never hit me with his or her fist.
11. My partner sometimes screams or yells at me when he or she is angry.
12. My partner has forced me to have sex when I didn't want to.
13. I have worried about my partner losing control of his or her anger.
14. My partner has never injured me physically.
15. I have never worried that my partner might become angry enough to hurt me.
16. My partner has left bruises or welts on my body.
17. My partner has never threatened to hurt me.
18. My partner has injured me so badly that I required medical treatment.
19. When angry, my partner sometimes threatens to leave me.

Table 2

Structure and Pattern Coefficients for Aggression Scale
When Factors are Uncorrelated

Item	Pattern Coefficients Factor I	Pattern Coefficients Factor II	Structure Coefficients Factor I	Structure Coefficients Factor II
1	.67	.34	.67	.34
2	.28	.71	.28	.71
3	.63	.46	.63	.46
4	.58	.48	.58	.48
5	.81	.26	.81	.26
6	.55	.40	.55	.40
7	.70	.25	.70	.25
8	.36	.67	.36	.67
9	.73	.23	.73	.23
10	.38	.56	.38	.56
11	.70	.10	.70	.10
12	.13	.64	.13	.64
13	.67	.40	.67	.40
14	.33	.75	.33	.75
15	.51	.59	.51	.59
16	.29	.80	.29	.80
17	.42	.68	.42	.68
18	-.04	.69	-.04	.69
19	.65	.05	.65	.05

Table 3

Structure and Pattern Coefficients for Aggression Scale
When Factors are Correlated

	Pattern Coefficients		Structure Coefficients	
Item	Factor I	Factor II	Factor I	Factor II
1	.65	.16	.74	.50
2	.11	.70	.46	.75
3	.57	.31	.73	.60
4	.51	.35	.69	.61
5	.83	.04	.85	.46
6	.50	.27	.64	.53
7	.71	.05	.74	.41
8	.20	.63	.53	.74
9	.75	.02	.76	.41
10	.27	.50	.52	.64
11	.76	-.11	.70	.27
12	-.04	.68	.30	.66
13	.64	.23	.76	.55
14	.15	.73	.52	.81
15	.40	.50	.65	.70
16	.09	.80	.50	.85
17	.27	.62	.59	.76
18	-.25	.78	.15	.66
19	.72	-.16	.64	.21

Table 4

Results of Literature Review

	Extraction			Rotation		Matrices Provided		Analysis	
Year	CFA	PC	NS	Varimax	NS	Yes	No	P	S
1980	29	29	43	100	0	71	29	57	43
1990-94	69	6	25	63	19	69	31	6	94

Note: Values are percentages. NS=not specified. P=primary focus of study. S=secondary focus of study.